



XX Workshop de Investigadores en Ciencias de la Computación
26 y 27 de abril de 2018 - Corrientes - Argentina

LIBRO DE ACTAS

Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste

Red de Universidades con Carreras de Informática (RedUNCI)



Dapozo, Gladys

XX Workshop de Investigadores en Ciencias de la Computación - WICC 2018 : libro de actas / Gladys Dapozo ; Patricia Pesado ; compilado por Gladys Dapozo ; Emanuel Irrazabal. - 1a ed compendiada. - Corrientes : Universidad Nacional del Nordeste. Facultad de Ciencias Exactas, 2018.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-3619-27-4

1. Bases de Datos. 2. Minería de Datos. 3. Ingeniería de Software. I. Pesado, Patricia II. Dapozo, Gladys , comp. III. Irrazabal, Emanuel , comp. IV. Título.

CDD 004.071

Indice de Artículos

Agentes y Sistemas Inteligentes	1
Desarrollo de Habilidades sociales y cognitivas a través de Sistemas Inteligentes	2
Operadores de Cambio de Creencias No-Priorizada en Sistemas Argumentativos Aplicado para Programación Lógica Rebatible	7
Aplicación de técnicas de Inteligencia Artificial para evaluar la producción científico-académica de investigadores de Universidades públicas del Noroeste Argentino	12
Diseño de Algoritmos Inteligentes Aplicados a Interfaces Humano-Computador e Internet de las Cosas	17
Monitoreo de carga por métodos no invasivos en el hogar argentino utilizando redes neuronales	22
Deep Neural Network para Análisis Acústico	27
Modelo basado en la Toma Decisiones con Criterios Múltiples para la elección de metodologías de Data Science	32
Diseño de Método de Ensamble Homogéneo para Clasificadores Débiles usando un esquema de reducción de datos simultaneo basado un enfoque co-evolutivo.	37
Aprendizaje Automático, Aplicaciones en Visión por Computadora	42
Redes Inteligentes de Agua: Factores y métodos para la predicción del consumo residencial de agua potable	46
Soluciones inteligentes para el desarrollo urbano sostenible	50
Metaheurísticas aplicadas de problemas de Scheduling con restricciones	56
Análisis automático de grandes volúmenes de datos en redes sociales mediante minería de textos combinado con algoritmos inteligentes	60
Integración de bases de creencias manteniendo coherencia y consistencia con criterio unificado de credibilidad	65
Optimización a gran escala usando metaheurísticas	70
Desarrollo de Sistemas de Scheduling para la Nube	75
Modelo de Inteligencia Artificial aplicados al Trading Algorítmico	80
GA2LA: Grupo de Aprendizaje Automático, Lenguajes y Autómatas	84
Sistemas de análisis textual en formato no estructurado	89
Procesamiento de textos estructurados	94
Arquitectura, Redes y Sistemas Operativos	98
Análisis de tráfico Multicast de video H264/Theora en Redes Wi-Fi IEEE 802.11ac	99
Sistemas de comunicación robustos para infraestructuras avanzadas de medición de energía implementadas sobre PLC (Power Line Communications)	104
Nuevo Modelo de Decisión para Gestión de Tráfico en Redes	109
Metodología para la selección de recursos computacionales gestionados con tecnologías de cloud computing en ambientes educativos	114
Nueva propuesta para la administración de recursos y procesos en sistemas distribuidos	118
Visualización Gráfica de resultados de Simulaciones de Redes de Sensores Inalámbricos	123

Implementación de sistemas de control automático para cultivos hidropónicos en invernaderos de la provincia de Misiones	128
El problema de las comunicaciones rurales: estudio y selección de las mejores soluciones	133
Análisis de Protocolos de Comunicaciones para Internet de las Cosas	138
Hub Of Things: Concentrador para el Internet de las Cosas	142
Análisis de Eficiencia y Tolerancia a Fallos en Arquitecturas Multiprocesador para Aplicaciones de Procesamiento de Datos	147
Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo	153
Contienda entre las variantes del protocolo TCP Vegas y Reno por los recursos de la red en un modelo híbrido simple	159
Evaluación de performance en Redes Definidas por Software	164
Servidor virtual para detección de intrusos y ataques en IPV6	169
Puerta de Enlace para Internet de las Cosas usando Computadora Industrial Abierta	174
Predicción y alerta temprana de incendios forestales mediante integración de WSN e imágenes satelitales en un método de reducción de incertidumbre guiado por datos	179
CIAA en la industria misionera: Fabrica de ladrillos. Integración con software SCADA y ERP	184
Rendimiento de Aplicaciones Web en plataformas de contenedores de código abierto	189
Bases de Datos y Minería de Datos	194
Modelo de Decisión para la Validación de Métodos de Imputación Mediante la Utilización de Algoritmos de Minería de Datos	195
Reconocimiento de Patrones Genéticos por medio de Grafos	200
Calidad Universitaria mediante Técnicas del Data Mining	205
Avances en selección de biomateriales utilizados en implantes dentales aplicando técnicas de minería de datos	209
Estructuras de Datos y Algoritmos Eficientes para Búsquedas Web y Procesamiento de Grandes Datos	214
Procesamiento Eficiente de Grafos Masivos para Aplicaciones en Redes Sociales	220
NoSQL: Modelos de datos y sistemas de gestión de bases de datos	225
Infraestructura de Acceso a Datos Primarios con Aporte de Semántica en Repositorios Digitales	229
Sistemas Recomendadores aplicados en Educación	234
Benchmarking de Bases de Datos NoSQL para el almacenamiento de Modelos Semánticos	238
Visualización en Ciencia de Datos	242
Análisis de similitud en documentos de texto mediante técnicas de ciencia de datos basadas en aprendizaje profundo (Deep Learning)	246
Construcción de una arquitectura BI para optimización de las decisiones a partir del estudio del nivel de madurez de los entornos tecnológicos empresariales	251
Recuperación de la Información	255
Estudio y análisis de técnicas de modelado de grandes volúmenes de datos jurídicos	261
Diseño y Construcción de Procesos de Explotación de Información para el Área de Ciencias de la Computación	265
Hacia la optimización del uso de datos abiertos en el ámbito público	270
Técnicas de unificación de datos para la visualización de grandes volúmenes de datos	275
Bases de Datos y Aplicación de Técnicas de Avanzadas	280
Aplicación de minera de datos para facilitar el tratamiento de las normas de producción de atributos semánticos en idioma español	285
Tecnologías de la Web Semántica aplicadas al tratamiento de documentos jurídicos electrónicos	290

AVANCES EN SELECCIÓN DE BIOMATERIALES UTILIZADOS EN IMPLANTES DENTALES APLICANDO TÉCNICAS DE MINERÍA DE DATOS

Ganz, Nancy B.^(a); Domínguez, Facundo A.^(b); Ares Alicia E.^(a); Kuna, Horacio D.^(b)

(a) Programa de Materiales y Físicoquímica; Laboratorio de Ciencia de los Materiales; Instituto de Materiales de Misiones; CONICET; Facultad de Ciencias Exactas Químicas y Naturales; Universidad Nacional de Misiones.

(b) Departamento de Informática; Facultad de Ciencias Exactas Químicas y Naturales; Universidad Nacional de Misiones.

nancy.bea.ganz@gmail.com

RESUMEN

En Argentina, como en otros países del mundo se fabrican biomateriales para diferentes aplicaciones con el objetivo de restaurar las funciones del cuerpo humano, como es el caso de los implantes dentales. Sin embargo, es necesario contar con información suficientemente calificada y accesible sobre: tipos de implantes dentales, propiedades de los biomateriales, causas de fallo, características y condiciones de salud de los pacientes que requieren de estos implantes. La carencia de un registro digital con datos reales de historias clínicas de pacientes que se han sometido a procesos quirúrgicos de colocación de implantes dentales, dificulta la tarea de análisis y extracción de conocimiento oculto para los especialistas implantólogos, sobre la relación implante dental - condición del paciente. De aquí, surge la necesidad de crear un registro automatizado que reúna las variables que representan el proceso de colocación de un implante dental, con el objetivo de identificar los factores que contribuyen al éxito o al fracaso de los implantes dentales colocados en la Provincia de Misiones, a través de la aplicación de técnicas de minería de datos, mediante el diseño de un procedimiento con métodos híbridos, y la aplicación de la metodología CRISP-DM para asentar el proceso.

Palabras Clave: *Minería de Datos, Biomateriales, Implantes Dentales.*

CONTEXTO

Esta línea de investigación se lleva a cabo dentro del Programa de Materiales y Físicoquímica (PROMyF) en el Laboratorio de Ciencia de los Materiales del Instituto de Materiales de Misiones (IMAM), de la Facultad en Ciencias Exactas, Químicas y Naturales (FCEQyN), de la Universidad Nacional de Misiones (UNaM), en el marco de un plan de tesis doctoral. Está financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) a través de una “Beca Interna Doctoral” otorgada por Resolución D N° 4869. Además, cuenta con el auspicio del Colegio de Odontólogos de la Provincia de Misiones.

1. INTRODUCCIÓN

La falta de un registro digital de implantes dentales con datos sobre condiciones del paciente, el proceso y el implante, hace dificultosa la tarea de investigación para extraer conocimiento que podría llegar a ayudar al proceso de oseointegración de un implante.

Los implantes dentales son biomateriales, es decir son materiales biológicos que reemplazan y/o restauran la función de un tejido o de un órgano vivo (pieza dental). Son inertes, debido a que están diseñados para ser implantado tanto dentro como fuera del cuerpo humano. [1]

El proceso de oseointegración define el éxito del implante, y éste depende de tres factores principales: propiedades y biocompatibilidad del implante, condiciones de salud del paciente, y la habilidad del cirujano. [2]

Sin embargo, la rehabilitación oral a través de implantes dentales puede presentar riesgos relacionados con la etapa del proceso de oseointegración implante/hueso. Estos riesgos pueden estar relacionados con las condiciones de salud del paciente, la técnica quirúrgica empleada por el especialista implantólogo, el tipo de implante, así como el tabaquismo. [3]

Actualmente, el campo de la minería de datos ha tenido muchos avances respecto a la aplicación y desarrollo de técnicas en el sector de la salud, para la predicción de enfermedades y para la toma de decisiones en base al análisis de grandes cantidades de datos. Por ejemplo: [4] Extrae datos de Historias Clínicas Electrónicas y utiliza la combinación de los algoritmos J48 y K-Means, para contribuir al diagnóstico de la hipertensión arterial. [5] Compara Árbol de Decisión, Clasificación Bayesiana, métodos predictivos como K-Nearest Neighbors (KNN) y Redes Neuronales para la predicción de enfermedad cardíaca. [6] Utiliza Redes Neuronales y Reglas de Asociación, para la detección y clasificación de tumores en mamografía digital. [7] Utiliza Redes Neuronales para discriminar grupos entre pacientes que se encontraban en terapia intensiva. [8] Utiliza Redes Bayesianas y Árboles de Decisión, además evalúa la utilidad de la metodología bayesiana en la predicción y el diagnóstico médico de enfermedades complejas. En [9] presenta la aplicación de métodos de aprendizaje automático (Red Neuronal, Máquina de Vector Soporte y Árbol de Decisión) al problema de predicción de caries en niños. En [10] presenta un estudio de minería de datos para formular nuevas hipótesis sobre la asociación entre la periodontitis y la diabetes tipo II.

Existe una gran cantidad de trabajos de minería de datos aplicados a la detección temprana y tratamiento de enfermedades oncológicas, así como en la identificación, prevención y

reducción de riesgos. Sin embargo, no se han encontrado trabajos que apliquen técnicas de minería de datos al campo de los biomateriales, específicamente a los implantes dentales, en [11] se muestra un análisis estadístico de Regresión Logística Múltiple para determinar los factores que influyen en el éxito de los implantes dentales. En [12] hace un estudio donde determina si existe relación entre los fracasos de los implantes dentales y las enfermedades sistémicas (concretamente sobre la osteoporosis, hipertensión, diabetes e hipotiroidismo), en una población de pacientes sometidos a cirugía de implantes dentales en el hospital San José de Santiago de Chile. En [13] presenta un estudio comparativo de tres técnicas de aprendizaje automático (Máquina de Vector Soporte (SVM), Máquina de Vector Soporte Ponderado y Redes Neuronales RBF constructivas (RBF-DDA) con parámetro de selección) para la predicción del éxito de los implantes dentales.

Si bien estos trabajos utilizan datasets de implantes dentales, no centran su atención al biomaterial, es por esto que esta línea de investigación busca abordar conjuntamente el estudio de las características del implante utilizado.

La detección de datos anómalos usando técnicas de Minería de Datos permite detectar grupos de datos que pueden ser de especial interés en temas relacionados como por ejemplo con el análisis de los biomateriales.[14]

Para documentar proyectos de Minería de Datos, existen procesos diseñados para guiar al usuario a través de una sucesión formal de pasos. Entre las metodologías más importantes encontramos: SEMMA, CRISP-DM [15] y P3TQ.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Es de vital importancia un registro digital con datos de historias clínicas sobre la colocación de implantes dentales, con variables que representen información relacionada al implante y a las condiciones de salud del paciente. En este contexto se reconoce como objetivo de interés para la presente línea de investigación, la creación de un registro digital de implantes dentales y el desarrollo de un procedimiento a través de una metodología híbrida, para facilitar la extracción de conocimiento útil para el especialista y actores de interés, e identificar los factores que contribuyan al éxito o al fracaso de los implantes dentales, así como las condiciones óptimas que debe tener el paciente y el implante dental.

Además, estudiar las propiedades mecánicas, químicas y físicas de los biomateriales utilizados en la implantología dental, para una mejor comprensión de la funcionalidad y resistencia de los implantes. Conjuntamente, lograr una taxonomía según el origen de fabricación y según el tipo de biomaterial utilizado en la industria del implante.

3. RESULTADOS OBTENIDOS/ESPERADOS

Entre los resultados obtenidos, destacamos la creación de un registro digital (datasets) con datos reales de historias clínicas de pacientes que se han sometido al proceso quirúrgico de colocación de implantes dentales en las localidades más características de la Provincia de Misiones.

Así mismo, se está trabajando en conjunto con expertos en el área de Minería de Datos, Biomateriales, Especialista en Implantología y Rehabilitación Compleja, Especialista en Patología Bucal e Implantología Oral, así como en Estadística.

Para comenzar con el proceso de extracción de conocimiento, los datos utilizados en este trabajo fueron obtenidos manualmente a través del diseño y corroboración de un formulario por parte de los expertos en el área. Así como, sucesivas entrevistas y charlas con más de 30

especialistas en implantología oral de toda la provincia de Misiones, con el fin de determinar cuáles son los datos de importancia a la hora de considerar la colocación de un implante dental.

Los datos del datasets se agrupan en 4 secciones:

- **Datos del Paciente:** referidos a los antecedentes y condiciones médicas de los pacientes a la hora de la intervención.
- **Datos del Implante:** referidos a las características del implante utilizado por el especialista implantólogo.
- **Datos de la Fase Quirúrgica:** referidos al procesamiento de intervención quirúrgica y mejoramiento del lecho óseo del paciente.
- **Datos del Seguimiento Postoperatorio:** referidos al resultado del proceso de colocación del implante, es decir si el proceso fue exitoso o fracaso.

Estas 4 secciones se encuentran alojadas en un solo archivo de datos, denominado BD_IMPLANTES.csv.

En base a la metodología CRISP-DM, la *fase 1 de comprensión del problema* abarco la definición del objetivo del presente trabajo, el cual es la determinación de los factores que contribuyen al éxito de los implantes dentales, así mismo se evaluó la situación de la carencia de un registro de estas características, se fijó el objetivo de minería de datos que es el diseño de un procedimiento a través de una metodología híbrida. Se precisó el plan de trabajo, que consistió en el relevamiento bibliográfico, investigación de la situación actual de la implantología en la provincia, definición de los expertos en el área, iniciar la recolección de datos y comenzar con la preparación de estos datos. En la *fase 2 de comprensión de los datos* se plasmó la recolección de los datos iniciales, donde se observa que los datos se agrupan en 4 secciones: datos del paciente, datos del implante, datos de la fase quirúrgica y datos del seguimiento postoperatorio. Así mismo, el

datasets cuenta con un total de 34 variables y 1050 casos de implantes colocados en la provincia de Misiones. Existen variables cuantitativas (como por ejemplo: *edad*, *longitud* y *diametro*) y variables cualitativas categóricas (como por ejemplo: *nivel_prof*, *pieza_dentaria*, *protoc_carga*, *seguir_postop*, entre otras). Seguidamente se detalla el contenido (columnas) del datasets “*BD_IMPLANTES.csv*” correspondiente a cada sección:

<i>Datos del Paciente</i>	edad, genero, ocupacion, o_s, enfermedad, fuma, alcoholismo, periodontitis, riesgo_periodontal, ingest_med, alergia.
<i>Datos del Implante</i>	material, marca, diseño, longitud, diametro, conexion, procedencia.
<i>Datos de la Fase Quirúrgica</i>	fecha_inter, lugar_proc, nivel_prof, pieza_dentaria, protoc_carga, exodoncia, expan_osea, elev_seno_maxilar, reg_tejidos_duros, reg_tejidos_blandos, tiempo_coloc, tipo_hueso,

	indic_protésica, comp_quirúrgica.
<i>Datos de Seguimiento Postoperatorio</i>	seguir_postop, observacion.

Actualmente, esta línea de investigación se encuentra en la etapa de preparación de los datos y verificación de la calidad de los mismos, mediante la aplicación de filtros.

Además, se busca la transformación y reducción de las variables que no aportaban ganancia de información al estudio de caso, con el fin de construir un datasets final para iniciar con el diseño de un procedimiento que utilice técnicas de minería de datos para resolver el problema planteado, e ir ajustando en función de las necesidades que se presenten.

Conjuntamente, se está profundizando el estudio y adaptabilidad del algoritmo máquina de vector soporte (SVM), esto en función de los trabajos citados que utilizan y recomiendan la aplicación de este método para datasets de este tipo.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación del “Programa de Materiales y Físicoquímica” de la FCEQyN – UNaM, con cuatro integrantes; de los cuales dos son docentes de la FCEQyN – UNaM (donde uno pertenece al programa mencionado y es investigador independiente del CONICET y el otro es investigador categoría I perteneciente al Departamento de Informática). El tercer integrante se encuentra realizando un doctorado; el cuarto y último integrante se encuentra desarrollando su tesina de grado de la carrera de Licenciatura en Sistemas de Información de la FCEQyN – UNaM.

5. BIBLIOGRAFÍA

- [1] B. D. Ratner, A. S. Hoffman, F. J. Schoen, and J. E. Lemons, *Biomaterials Science. An introduction to Materials in*

- Medicine. San Diego, California - USA: Elsevier Academic Press, 1996.
- [2] G. P. Kothiyal and A. Srinivasan, "Trends in Biomaterials", *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2016.
- [3] P. Laine, A. Salo, R. Kontio, S. Ylijoki, C. Lindqvist, and R. Suuronen, "Failed dental implants – clinical, radiological and bacteriological findings in 17 patients", *J. Cranio-Maxillofacial Surg.*, vol. 33, no. 3, pp. 212–217, Jun. 2005.
- [4] F. Dávila Hernández and Y. Sánchez Corales, "Técnicas de minería de datos aplicadas al diagnóstico de entidades clínicas", *Rev. Cuba. Informática Médica*, vol. 4, no. 2, pp. 174–183, 2012.
- [5] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", *Int. J. Comput. Appl.*, vol. 17, no. 8, pp. 43–48, 2011.
- [6] M. Antonie, A. Coman, and O. R. Zaiane, "Application of Data Mining Techniques for Medical Image Classification", *Proc. Second Int. Work. Multimida Data Min.*, pp. 94–101, 2001.
- [7] C. Arizmendi, E. Romero, R. Alquezar, P. Caminal, I. Díaz, S. Benito, and B. Giraldo, "Data mining of patients on weaning trials from mechanical ventilation using cluster analysis and neural networks", *2009 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 4343–4346, 2009.
- [8] G. Solarte and Y. Castro, "Modelo híbrido para el diagnóstico de enfermedades cardiovasculares basado en inteligencia artificial", *Rev. Tecnura*, vol. 16, no. 33, pp. 35–52, 2012.
- [9] R. D. Montenegro, A. L. I. Oliveira, G. Cabral, C. R. T. Katz, and A. Rosenblatt, "A comparative study of machine learning techniques for caries prediction", *20th IEEE Int. Conf. Tools with Artif. Intell.*, vol. 2, pp. 477–481, 2008.
- [10] U. Covani, S. Marconcini, G. Derchi, A. Barone, and L. Giacomelli, "Relationship between human periodontitis and type 2 diabetes at a genomic level: a data-mining study", *J. Periodontol.*, vol. 80, no. 8, pp. 1265–1273, 2009.
- [11] J. E. B. Tamez, F. N. Zilli, L. A. Fandiño, and J. M. Guizar, "Factores relacionados con el éxito o el fracaso de los implantes dentales colocados en la especialidad de Prosthodontia e Implantología en la Universidad de La Salle Bajío", *Revi. Esp. Cirugía Oral y Maxilofac.*, pp. 1–9, 2016.
- [12] J. Domínguez, J. Acuña, M. Rojas, J. Bahamondes, and S. Matus, "Study of association between systemic diseases and dental implant failure", *Rev. Clínica Periodoncia, Implantol. y Rehabil. Oral*, vol. 6, no. 1, pp. 9–13, 2013.
- [13] A. L. I. Oliveira, C. Baldisserotto, and J. Baldisserotto, "A comparative study on machine learning techniques for prediction of success of dental implants", *Mex. Int. Conf. Artif. Intell.*, vol. 3789 LNAI, pp. 939–948, 2005.
- [14] H. Kuna, R. G. Martinez, and F. Villatoro, "Automatic Outliers Fields Detection in Databases", *J. Model. Simul. Syst. HyperSciences Publ.*, vol. 3, no. 1, p. 14–20., 2012.
- [15] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer and R. Wirth, *CRISP-DM 1.0: Step-by-step data mining guide*, 2000.